10-11-2002

Long-Step Homogeneous Interior-Point Method for P*-Nonlinear Complementarity Problem

Goran Lesaja
Georgia Southern University, goran@georgiasouthern.edu

Yugoslav Journal of Operations Research
12 (2002), Number 1, 17-48

LONG-STEP HOMOGENEOUS INTERIOR-POINT ALGORITHM FOR THE P_*-NONLINEAR COMPLEMENTARITY PROBLEMS*

Goran LEŠAJA

Department of Mathematics and Computer Science
Georgia Southern University
Statesboro, USA

Abstract: A P_*-Nonlinear Complementarity Problem as a generalization of the P_*-Linear Complementarity Problem is considered. We show that the long-step version of the homogeneous self-dual interior-point algorithm could be used to solve such a problem. The algorithm achieves linear global convergence and quadratic local convergence under the following assumptions: the function satisfies a modified scaled Lipschitz condition, the problem has a strictly complementary solution, and certain submatrix of the Jacobian is nonsingular on some compact set.

Keywords: P_*-nonlinear complementarity problem, homogeneous interior-point algorithm, wide neighborhood of the central path, polynomial complexity, quadratic convergence.

1. INTRODUCTION

The nonlinear complementarity problem (NCP), as described in the next section, is a framework which can be applied to many important mathematical programming problems. The Karush-Kuhn-Tucker (KKT) system for the convex optimization problems is a monotone NCP. Also, the variational inequality problem can be formulated as a mixed NCP (see Farris and Pang [6]). The linear complementarity problem (LCP), a special case of NCP, has been studied extensively. For a comprehensive treatment of LCP see the monograph of Cottle et al. [4].

---

* Some results contained in this paper were first published in the author's Ph.D. thesis. Further research on this topic was supported in part by Georgia Southern Faculty Research Subcommittee Faculty Research Grant. AMS subject classification: 90C05, 65K05.

The interior-point methods, originally developed for the linear programming problem (LP), have been successfully extended to LCP, NCP, and the semidefinite programming problems (SDP). A number of papers dealing with LP and LCP, is extensive. Many topics, like the existence of the central path, global and local convergence and implementation issues have been studied extensively. Fewer papers are devoted to NCP. Among the earliest are the important works of Dikin [5], McLinden [19], and Nesterov and Nemirovskii [24].

In the series of papers Kojima et al. [14, 15, 13, 16, 17, 11] studied different classes of NCP when the function was P_0-function, uniform P-function, or monotone function. They analyzed the central paths of these problems and proposed the continuation, or interior-point methods to solve them. No polynomial global and/or local convergence result were given.

A number of other interior-point algorithms for monotone NCP has been developed, among them Potra and Ye [30], Andersen and Ye [1], Guller [7], Nesterov [23], Monteiro et al. [21], Sun and Zhao [31], Tseng [33, 32], Wright and Ralph [35]. Polynomial global convergence for many of the algorithms has been proven when the function is monotone and satisfies certain smoothness condition. The most general one is a self-concordant condition of Nesterov and Nemirovskii [24]. Other conditions include the relative Lipschitz condition of Jarre [9], and the scaled Lipschitz condition of Potra and Ye [30].

In the linear case, that is for LCP, the above mentioned smoothness conditions are unnecessary to prove polynomial global and local convergence of the various interior-point methods. Moreover, the convergence results have been proven for more general classes of functions than monotone functions. Among others is a P_*-LCP introduced by Kojima et al. [12]. See also Miao [20], Ji et al. [10], Potra and Sheng [28], Anitescu et al. [3, 2].

In this paper we study the P_*-NCP that generalizes monotone NCP in the similar way in which P_*-LCP generalizes monotone LCP. This class was introduced independently by the authors [18] and Jansen et al. [8]. There are few papers that study the class of P_*-NCP. Recently Peng et al. [26] analyzed interior-point method for P_*-NCP using self-regular proximities that they initially introduced for LP and LCP. In Jansen et al. [8] the definition of the P_*-functions is indirect, it is based on the P_*-property of the Jacobian matrix, while our definition deals directly with the function. We also provide the equivalency proof between the two definitions (Lemma 2.1). A similar approach is adopted by Peng et al. [26].

The second objective of the paper is to prove linear global and quadratic convergence of the interior-point method for the P_*-NCP. We use a long-step version of the homogeneous, self-dual, interior-point algorithm of [1]. In [1] polynomial global convergence of the short-step version of the algorithm was analyzed but no local convergence result was established. Based on the analysis in [31] and [37], we prove that iteration sequence converges to the strictly complementary solution with R-order at least 2, while primal-dual gap converges to zero with R-order and Q-order at least 2 under the following list of assumptions described later in the text: the existence of a strictly complementary solution (ESCS), the modified scaled Lipschitz condition of

Potra and Ye (SLC), and the nonsingularity of the Jacobian submatrix (NJS). This set of assumptions is weaker than the one in [31]. We show that Assumption 3 in [31] is a consequence of the scaled Lipschitz condition (Lemma 5.6).

One more comment is in order. Since most of the smoothness conditions were introduced for monotone functions, we have chosen to modify the scaled Lipschitz condition of Potra and Ye [30] to be able to handle P_*-functions. For the same purpose in [8] a different modification of scaled Lipschitz condition has been introduced (Condition 3.2) and its relation to some known conditions has been discussed. On the other hand, Peng et al. [26] used a generalization of Jarre's relative Lipschitz condition.

The paper is organized as follows: In Section 2 we formulate P_*-NCP. In Section 3 we discuss a homogeneous model for P_*-NCP and introduce a long-step infeasible interior-point algorithm for this model. Global convergence is analyzed in Section 4. We end the paper with analysis of a local convergence contained in Section 5.

2. PROBLEM

We consider a nonlinear complementarity problem (NCP) of the form

$$(NCP) \quad s = f(x), \ x \geq 0, \ x^T s = 0 ,$$

where $x, s \in R^n$ and f is a C^1 function $f : R_+^n \to R^n$.

Denote a feasible set of NCP by

$$\mathcal{F} = \{(x,s) \in R_+^{2n} : s = f(x)\} ,$$

and its solution set by

$$\mathcal{F}^* = \{(x^*, s^*) \in \mathcal{F} : x^{*T} s^* = 0\} .$$

For any given $\varepsilon > 0$ we define the set of ε-approximate solutions of NCP as

$$\mathcal{F}_\varepsilon = \{(x,s) \in R_+^{2n} : x^T s < \varepsilon, \ \| s - f(x) \| < \varepsilon\} .$$

If f is a linear function

$$f(x) = Mx + q ,$$

where $M \in R^{n \times n}$ and $q \in R^n$, then the problem reduces to LCP. The LCP has been studied for many different classes of matrices M (see [4, 12]). We list some:

- Skew-symmetric matrices (SS):

$$(\forall x \in R^n)(x^T M x = 0) . \tag{2.1}$$

- Positive semidefinite matrices (PSD):

$$(\forall x \in R^n)(x^T M x \geq 0) . \tag{2.2}$$

- P -matrices: Matrices with all principal minors positive or equivalently:

$$(\forall x \in R^n,\ x \neq 0)\,(\exists i \in I)(x_i(Mx)_i > 0)\,. \tag{2.3}$$

- P_0 -matrices: Matrices with all principal minors nonnegative or equivalently

$$(\forall x \in R^n,\ x \neq 0)\,(\exists i \in I)(x_i \neq 0 \text{ and } x_i(Mx)_i \geq 0)\,. \tag{2.4}$$

- Sufficient matrices (SU): Matrices which are column and row sufficient
  - Column sufficient matrices (CSU)

$$(\forall x \in R^n)\,(\forall i \in I)(x_i(Mx)_i \leq 0 \Rightarrow x_i(Mx)_i = 0)\,. \tag{2.5}$$

  - Row sufficient matrices (RSU): M is row sufficient if M^T is column sufficient.
- $P_*(\kappa)$: Matrices such that

$$(1+4\kappa)\sum_{i \in \mathfrak{T}^+(x)} x_i(Mx)_i + \sum_{i \in \mathfrak{T}^-(x)} x_i(Mx)_i \geq 0, \quad \forall x \in R^n$$

where

$$\mathfrak{T}^+(x) = \{i : x_i(Mx)_i > 0\},\ \mathfrak{T}^-(x) = \{i : x_i(Mx)_i < 0\}\,,$$

or equivalently

$$x^T Mx \geq -4\kappa \sum_{i \in \mathfrak{T}^+(x)} x_i(Mx)_i, \quad \forall x \in R^n\,, \tag{2.6}$$

and

$$P_* = \bigcup_{\kappa \geq 0} P_*(\kappa)\,. \tag{2.7}$$

The relationship between some of the above classes is as follows

$$SS \subset PSD \subset P_* = SU \subset CS \subset P_0, \quad P \subset P_*, \quad P \cap SS = \varnothing\,. \tag{2.8}$$

Some of these relations are obvious, like $PSD = P_*(0) \subset P_*$ or $P \subset P_*$, while others require a proof which can be found in [12, 4, 34].

The above classes can be generalized for nonlinear functions as follows:

- Monotone functions

$$(\forall x^1, x^2 \in R^n)\,((x^1 - x^2)^T (f(x^1) - f(x^2)) \geq 0)\,, \tag{2.9}$$

are a generalization of positive semidefinite matrices (PSD).

- P -functions

$$(\forall x^1, x^2 \in R^n, x^1 \neq x^2)\,(\exists i \in I)\,((x_i^1 - x_i^2)(f_i(x^1) - f_i(x^2)) > 0)\,, \tag{2.10}$$

are a generalization of P -matrices. A special case of P -function is uniform P -function with parameter $\gamma > 0$

$$(\forall x^1, x^2 \in R^n, x^1 \neq x^2)\,(\exists i \in I)\,((x_i^1 - x_i^2)(f_i(x^1) - f_i(x^2)) \geq \gamma \,\| x^1 - x^2 \|^2)\,. \tag{2.11}$$

- P_0 -functions

$$(\forall x^1, x^2 \in R^n, x^1 \neq x^2)\,(\exists i \in I)\,(x_i^1 - x_i^2 \neq 0, (x_i^1 - x_i^2)(f_i(x^1) - f_i(x^2)) \geq 0), \qquad (2.12)$$

are a generalization of P_0 -matrices.

Below we give a definition of $P_*(\kappa)$ -functions generalizing the definition of $P_*(\kappa)$ -matrices.

- $P_*(\kappa)$ -functions

A function f belongs to the class of $P_*(\kappa)$ -functions if for each $x^1, x^2 \in R^n$ the following inequality holds

$$(x^2 - x^1)^T (f(x^1) - f(x^2)) \geq -4\kappa \sum_{i \in \mathfrak{T}_f^+} (x_i^2 - x_i^1)(f_i(x^1) - f_i(x^2)),$$

where

$$\mathfrak{T}_f^+ = \{i \in \{1,...,n\} : (x_i^2 - x_i^1)(f_i(x^1) - f_i(x^2)) > 0\},$$

and $\kappa \geq 0$ is a constant.

- P_* -functions

A function f is a P_* -function if there exists $\kappa \geq 0$ such that f is a $P_*(\kappa)$ -function. This is equivalent to

$$P_* = \bigcup_{\kappa \geq 0} P_*(\kappa).$$

The classes of $P_*(\kappa)$ -functions and P_* -functions were introduced independently in Jansen et al. [8] and first in the author's Ph. D. thesis [18]. Note that the class of monotone functions, considered in the most papers about NCP, is included as a special case for $\kappa = 0$, i.e. as $P_*(0)$ case. Throughout the paper we assume that the function f is a P_* -function.

The following lemma establishes a relationship between $P_*(\kappa)$ -property of the function f and its Jacobian matrix ∇f.

Lemma 2.1. *The function f is a $P_*(\kappa)$ -function iff ∇f is a $P_*(\kappa)$ -matrix.*

Proof: Suppose first that f is a $P_*(\kappa)$ -function,

$$(x^2 - x^1)^T (f(x^1) - f(x^2)) \geq -4\kappa \sum_{i \in \mathfrak{T}^+} (x_i^2 - x_i^1)(f_i(x^1) - f_i(x^2)).$$

Since f is a C^1 function, the following equations hold

$$f(x+h) - f(x) = \nabla f(x) h + o(h),$$

$$f_i(x+h) - f_i(x) = \sum_{j=1}^{n} (\nabla f(x))_{ij} h_j + o(h).$$

If we denote $h = x^2 - x^1$ and if we use the above equations, then the left hand side of the above inequality becomes

$$\begin{aligned}(x^2 - x^1)^T (f(x^1) - f(x^2)) &= h^T (f(x+h) - f(x)) \\ &= h^T \nabla f(x) h + o(h^2),\end{aligned}$$

while the right hand side can be written as

$$\begin{aligned}-4\kappa \sum_{i \in \mathfrak{T}^+} (x_i^2 - x_i^1)(f_i(x^i) - f_i(x^2)) &= -4\kappa \sum_{i \in \mathfrak{T}^+} h_i (f_i(x+h) - f_i(x)) \\ &= -4\kappa \sum_{i \in \mathfrak{T}^+} \sum_{j=1}^{n} (\nabla f(x))_{ij} h_j h_i + o(h^2) \\ &= -4\kappa \sum_{i \in \mathfrak{T}^+} h_i (\nabla f(x) h)_i + o(h^2).\end{aligned}$$

We get

$$h^T \nabla f(x) h \geq -4\kappa \sum_{i \in \mathfrak{T}^+} h_i (\nabla f(x) h)_i + o(h^2) .$$

Given u take $h = \varepsilon u$. The above inequality transforms to

$$\varepsilon^2 u^T \nabla f(x) u \geq -\varepsilon^2 4\kappa \sum_{i \in \mathfrak{T}^+} u_i (\nabla f(x) u)_i + o(\varepsilon^2) .$$

Dividing the above inequality by ε^2 and taking the limit as $\varepsilon \to 0$ we have

$$h^T \nabla f(x) h \geq -4\kappa \sum_{i \in \mathfrak{T}^+} h_i (\nabla f(x) h)_i .$$

Hence $\nabla f(x)$ is a $P_*(\kappa)$ - matrix.

To prove the other implication, suppose that $\nabla f(x)$ is a $P_*(\kappa)$ -matrix, i.e., the above inequality holds. Using the mean value theorem for the function f we have

$$\begin{aligned}h^T (f(x+h) - f(x)) &= h^T \int_0^1 \nabla f(x+th) h \, dt \\ &= \int_0^1 h^T \nabla f(x+th) h \, dt \\ &\geq \int_0^1 \left(-4\kappa \sum_{i \in \mathfrak{T}^+} h_i (\nabla f(x+th) h)_i \right) dt \\ &= -4\kappa \sum_{i \in \mathfrak{T}^+} h_i \int_0^1 (\nabla f(x+th) h)_i \, dt \\ &= -4\kappa \sum_{i \in \mathfrak{T}^+} h_i (f_i(x+h) - f_i(x)).\end{aligned}$$

Hence $\nabla f(x)$ is a $P_*(\kappa)$ - function.

In [22] it was shown that the existence of a strictly complementary solution is necessary and sufficient to prove quadratic local convergence of an interior-point algorithm for the monotone LCP (see also [37]). This implies that we need to make the same assumption for the P_*-NCP.

Existence of a strict complementary solution (ESCS)

NCP has a strictly complementary solution, i.e., there exists a point $(x,s)\in \mathcal{F}^*$ such that

$$x+s>0 .$$

Unfortunately, even in the case of the monotone NCP the above assumptions are not sufficient to prove linear global and quadratic local convergence of the interior-point algorithm, thus additional assumptions are necessary. Therefore, additional assumptions are necessary for P_*-NCP as well. They will be introduced as they are needed later in the text.

3. ALGORITHM

In the development of the interior-point methods we can indicate two main approaches. The first is the application of the interior-point method to the original problem. In this case it is sometimes hard to deal with issues such as finding a feasible starting point detecting infeasibility or, more generally, determining nonexistence of the solution (it is known that monotone NCP may be feasible but still may not have a solution, which is not the case for the monotone LCP). Numerous procedures have been developed to overcome this difficulty ("big M" method, phase I - phase II methods, etc.). but none of them was completely satisfactory. It has been shown that a successful way to handle the problem is to build an augmented homogeneous self-dual model which is always feasible and then apply the interior-point method to that model. The "price" to pay is not that high (the dimension of the problem increases only by one) while on the other side benefits are numerous and important (the analysis is simplified, the size of the initial point or solutions is irrelevant due to the homogeneity, detection of infeasibility is solved in a natural way, etc.) This second approach originated in [38], and was successfully extended to LCP in [36], monotone NCP in [1], and SDP in [29].

Motivated by the above discussion in this paper we consider the augmented homogeneous self-dual model of [1] to accompany the original NCP.

$$(HNCP) \qquad \begin{aligned} & s=\tau f(x/\tau), \\ & \sigma=-x^T f(x/\tau), \\ & x^T s+\tau\sigma=0, \\ & (x,\tau,s,\sigma)\geq 0. \end{aligned}$$

Lemma 3.1. *HNCP is feasible and every feasible point is a solution point.*

The solutions of HNCP is related to the solutions of the original NCP as follows.

Lemma 3.2.

(i) *If* $(x^*, \tau^*, s^*, \sigma^*)$ *is a solution for HNCP and* $\tau^* > 0$, *then* $(x^*/\tau^*, s^*/\sigma^*)$ *is a solution for NCP.*

(ii) *If* (x^*, s^*) *is a solution for NCP, then* $(x^*, 1, s^*, 0)$ *is a solution for HNCP.*

The immediate consequence of the above lemma is the existence of a strict complementary solution for HNCP with $\tau^* > 0$ since in the previous section we assumed the existence of a strict complementary solution for NCP.

Using the first two equations in HNCP we can define an augmented transformation

$$\psi(x,\tau) = \begin{pmatrix} \tau f(x/\tau) \\ -x^T f(x/\tau) \end{pmatrix} : R_{++}^{n+1} \to R^{n+1} . \tag{3.1}$$

The augmented transformation has several important properties stated in the following lemma.

Lemma 3.3.

(i) ψ *is a* C^1 *homogeneous function with degree 1 satisfying*

$$(x,\tau)^T \psi(x,\tau) = 0 . \tag{3.2}$$

(ii) *The Jacobian matrix* $\nabla\psi(x,\tau)$ *of the augmented transformation (3.1) is given by*

$$\nabla\psi(x,\tau) = \begin{bmatrix} \nabla f(x/\tau) & f(x/\tau) - \nabla f(x/\tau)(x/\tau) \\ -f(x/\tau)^T - (x/\tau)^T \nabla f(x/\tau)^T & (x/\tau)^T \nabla f(x/\tau)(x/\tau) \end{bmatrix} \tag{3.3}$$

and following equality holds

$$(x/\tau)^T \nabla\psi(x/\tau) = -\psi(x/\tau)^T . \tag{3.4}$$

The proofs of the Lemma 3.1-3.3 can be found in [1]. Now we prove that if the augmented transformation ψ is a $P_*(\kappa)$-function then f is a $P_*(\kappa)$-function too.

Lemma 3.4. *If* ψ *is a* $P_*(\kappa)$*-function, then* f *is also a* $P_*(\kappa)$*-function.*

Proof: Using Lemma 2.1 we conclude that $\nabla\psi$ is $P_*(\kappa)$-matrix. From (3.3) and the fact that every principal submatrix of $P_*(\kappa)$-matrix is also a $P_*(\kappa)$-matrix (see [12]), it follows that ∇f is a $P_*(\kappa)$-matrix. Using again Lemma 2.1 we conclude that f is a $P_*(\kappa)$-function. ♦

It would be very desirable if the reverse implication is true as it is the case for monotone NCP. Unfortunately, that is not generally the case even for $P_*(\kappa)$-LCPs as shown by Peng et al. [25]. Thus, in what follows we will assume that ψ is a $P_*(\kappa)$-function.

Note that not all of the nice properties of the homogeneous model for monotone NCP could have been preserved for $P_*(\kappa)$ NCP. However, the homogeneous model still has a merit primarily because of its feasibility. In addition, the analysis that we provide in this paper holds if an interior-point method is used on the original problem rather than on the augmented homogeneous model.

The objective is to find ε -approximate solution of HNCP. We will do so by using a long-step primal-dual infeasible-interior-point algorithm. To simplify the analysis in the remainder of this paper we let

$$x = \begin{pmatrix} x \\ \tau \end{pmatrix}, \quad s = \begin{pmatrix} s \\ \sigma \end{pmatrix}. \tag{3.5}$$

A long-step algorithm produces the iterates $(x^k, s^k) \in R_{++}^{2n+2}$ belonging to

$$\mathcal{N}_\infty^-(\beta) = \left\{ (x,s) \geq 0 : Xs \geq \beta\mu e,\ \mu = \frac{x^T s}{n+1} \right\}, \quad 0 < \beta < 1 ,$$

which is the widest neighborhood of the central path

$$\boldsymbol{C}(t) = \{(x,s) \geq 0 : Xs = te,\ s - \psi(x) = tr^0\}, \quad 0 < t \leq 1 ,$$

where $(x^0, s^0) > 0$ is an initial point on the central path, r denotes a residual of the point (x,s)

$$r = s - \psi(x) , \tag{3.6}$$

so that $r^0 = s^0 - \psi(x^0)$, and X denotes a diagonal matrix corresponding to the vector x . If $\beta = 0$, then $\mathcal{N}_\infty^-(\beta)$ is the entire nonnegative orthant, and if $\beta = 1$, then $\mathcal{N}_\infty^-(\beta)$ shrinks to the central path $\boldsymbol{C}$.

Now we state the algorithm.

Algorithm 3.5

I (*Initialization*)

Let $\varepsilon > 0$ be a given tolerance, and let $\beta, \eta, \gamma \in (0,1)$ be the given constants. Suppose a starting point $(x^0, s^0) \in \mathcal{N}_\infty^-(\beta)$ is available. Calculate $\mu_0 = (x^0)^T s^0 / (n+1)$ and set $k = 0$.

S (*Step*)

Given $(x^k, s^k) \in \mathcal{N}_\infty^-(\beta)$ *solve the system*

$$\nabla\psi(x^k)\Delta x - \Delta s = \eta r^k , \tag{3.7}$$

$$S^k \Delta x + X^k \Delta s = \gamma\mu_k e - X^k s^k . \tag{3.8}$$

Let

$$x(\theta) = x^k + \theta\Delta x, \quad s(\theta) = \psi(x(\theta)) + (1 - \eta\theta) r^k , \tag{3.9}$$

and perform a line search to determine the maximal stepsize $0<\theta_k<1$ *such that*

$$(x(\theta_k),s(\theta_k))\in \mathcal{N}_\infty^-(\beta) \tag{3.10}$$

and $\mu(\theta_k)$ *minimizes* $\mu(\theta)$. *Set*

$$x^{k+1}=x(\theta_k),\quad s^{k+1}=s(\theta_k)\,. \tag{3.11}$$

T (*Termination*)

If

$$(x^{k+1},s^{k+1})\in \Psi_\varepsilon=\{(x,s)\ge 0: x^T s\le\varepsilon, \| s-\psi(x)\|\le\varepsilon\}\,, \tag{3.12}$$

then stop, otherwise set $k:=k+1$ *and go to (S).*

In the next two sections we will prove that there exist the values of the parameters for which the algorithm has polynomial global convergence and quadratic local convergence, provided that some additional assumptions, stated later in the text, are satisfied.

Now we give some basic properties of the direction $(\Delta x,\Delta s)$ and update $(x(\theta),s(\theta))$ calculated in the Algorithm 3.5.

Lemma 3.6. *Let* $(\Delta x,\Delta s)$ *be a solution of the system (3.7)-(3.8). Then*

$$(\Delta x)^T\Delta s=(\Delta x)^T\Delta\psi(x^k)\Delta x+\eta(1-\eta-\gamma)(n+1)\mu_k\,.$$

The proof of the above lemma can be found in [1].

The update (3.9) for $s(\theta)$ is obtained by approximating the residual $r=s-\psi(x)$ with its first order Taylor polynomial

$$s(\theta)-\psi(x(\theta))\approx s^k-\psi(x^k)+\theta(\Delta s-\nabla\psi(x^k)\Delta x)\,, \tag{3.13}$$

or by virtue of (3.7)

$$s(\theta)\approx\psi(x(\theta))+r^k-\theta\eta r^k\,.$$

Thus we set

$$s(\theta):=\psi(x(\theta))+(1-\theta\eta)r^k\,,$$

as stated in (3.9). Using (3.13) we have

$$\begin{aligned}X(\theta)s(\theta)&=X(\theta)(s^k+\theta\Delta s+\psi(x(\theta))-\psi(x^k)-\theta\nabla\psi(x^k)\Delta x)\\&=(X^k+\theta\Delta X)(s^k+\theta\Delta s)+X(\theta)(\psi(x(\theta))-\psi(x^k)-\theta\nabla\psi(x^k)\Delta x)\\&=X^k s^k+\theta(S^k\Delta x+X^k\Delta s)+\theta^2\Delta X\Delta s+(X^k+\theta\Delta X)(\psi(x(\theta))-\psi(x^k)-\theta\nabla\psi(x^k)\Delta x).\end{aligned}$$

If we denote the second order term in the above expression by

$$\begin{aligned} h(\theta) &= X^k s^k + \theta(S^k \Delta x + X^k \Delta s) + \theta^2 \Delta X \Delta s \\ &\quad + (X^k + \theta \Delta X)(\psi(x(\theta)) - \psi(x^k) - \theta \nabla \psi(x^k) \Delta x), \end{aligned} \tag{3.14}$$

then by virtue of (3.8) we obtain

$$X(\theta)s(\theta) = (1-\theta)X^k s^k + \theta\gamma\mu_k e + h(\theta). \tag{3.15}$$

Now the following lemma can easily be proved.

Lemma 3.7. *Consider the update* $(x(\theta), s(\theta))$ *given by (3.9). Then*

(i) $r(\theta) = (1-\theta\eta)r^k$,

(ii) $\mu(\theta) = (1-\theta(1-\gamma) + \theta^2\eta(1-\eta-\gamma))\mu_k$.

4. GLOBAL CONVERGENCE

In this section we prove polynomial global convergence of the Algorithm 3.5. If the function f is linear, i.e. if we have LCP, global convergence has been proven without any additional assumptions when f belongs to the P_*-class [20, 28, 10, 3]. This is not the case for f nonlinear. Global convergence has been proven for the monotone nonlinear function f under certain smoothness condition. The most general one is a self-concordant condition of Nesterov and Nemirovskii [24]. Other conditions include the relative Lipschitz condition of Jarre [9] and the scaled Lipschitz condition of Potra and Ye [30].

We adopt the following modification of the scaled Lipschitz condition.

Scaled Lipschitz condition (SLC)

There exists a monotone increasing function $v(\alpha):(0,1)\to(1,\infty)$ such that

$$\| X(f(x+\Delta x) - f(x) - \nabla f(x)\nabla x) \|_\infty \le v(\alpha) \,|\, \Delta x^T \nabla f(x) \Delta x \,|$$

whenever

$$\Delta x \in R^n, \quad x \in R^n_{++} := \{x \in R^n : x > 0\}, \quad \| X^{-1}\Delta x \|_\infty \le \alpha < 1. \qquad \blacklozenge$$

Other types of SLC have been used in the literature [1, 30, 31] with either ℓ_1 or ℓ_2 norm instead of ℓ_∞, and the constant has been used instead of the function v. Also the absolute value on the right-hand-side was not necessary because SLC was used for monotone functions for which $\Delta x^T \nabla f(x) \Delta x \ge 0$.

In [8] SLC was replaced with the new smoothness condition (Condition 3.2) to enable handling of the nonmonotone functions. Basically, under certain assumptions, Condition 3.2 requires the following inequality to hold

$$\| D(f(x+\Delta x) - f(x) - \nabla f(x)\Delta x) \| \le L \,\| D\nabla f(x)\Delta x \|,$$

where D is a certain diagonal matrix and L is a constant. The new condition essentially bounds the norm of the scaled second order remainder of the Taylor expansion of the function f by the norm of the first order term in that expansion, while SLC bounds it by the norm of the second order term. A condition similar to Condition 3.2 was recently introduced in [26] (Condition A.3).

The following lemma establishes the relation between SLC of the original function f and the augmented function ψ. Its proof is a trivial modification of the corresponding proof in [1].

Lemma 4.1. *If f satisfies SLC with $v = v_f$, then ψ satisfies SLC with*

$$v = v_\psi(\alpha) = \left(1 + \frac{2v_f(2\alpha/(1+\alpha))}{1-\alpha}\right)\left(\frac{1}{1-\alpha}\right).$$

To simplify the analysis, in what follows we assume that

$$\eta = 1 - \gamma. \tag{4.1}$$

Then from Lemma 3.6 and Lemma 3.7 we obtain

$$(\Delta x)^T \Delta s = (\Delta x)^T \nabla \psi(x^k) \Delta x, \tag{4.2}$$

$$\mu(\theta) = \frac{x(\theta)^T s(\theta)}{n+1} = (1-\eta\theta)\frac{(x^k)^T s^k}{n+1} = (1-\eta\theta)\mu_k, \tag{4.3}$$

$$r(\theta) = (1-\eta\theta) r^k, \tag{4.4}$$

which means that the infeasibility residual and the complementarity gap are reduced at the exactly same rate. The immediate consequence of (4.3) and (4.4) is that the issue of proving polynomial global convergence reduces to the problem of finding a positive lower bound $\hat{\theta}$ for the stepsize θ_k in the Algorithm 3.5 such that

$$\eta\hat{\theta} = \frac{C}{n^q},$$

where q is a rational number and C is a constant. For long-step algorithms (neighborhood $\mathcal{N}_\infty^-(\beta)$) the best possible q is $q = 1$, while for short-step algorithms (neighborhoods $\mathcal{N}_2(\beta)$) q can be reduced to $q = 1/2$.

We start the analysis by considering the main requirement in the algorithm 3.5 and that is, given the iterate $(x^k, s^k) \in \mathcal{N}_\infty^-(\beta)$, the new iterate $(x(\theta), s(\theta))$ must also belong to $\mathcal{N}_\infty^-(\beta)$. Using (3.15) and (4.3) we have

$$\begin{aligned} X(\theta)s(\theta) - \beta\mu(\theta)e &= (1-\theta)X^k s^k + \theta\gamma\mu_k e + h(\theta) - \beta(1-\eta\theta)\mu_k e \\ &\geq (1-\theta)\beta\mu_k e + \theta\gamma\mu_k e + h(\theta) - \beta(1-\eta\theta)\mu_k e \\ &\geq (\beta(1-\theta-1+\eta\theta) + \theta\gamma)\mu_k e - \| h(\theta) \|_\infty e \\ &= (1-\beta)\gamma\theta\mu_k e - \| h(\theta) \|_\infty e. \end{aligned} \tag{4.5}$$

Hence, if

$$\| h(\theta) \|_\infty \le (1-\beta)\gamma\theta\mu_k ,$$

then

$$(x(\theta), s(\theta)) \in \mathcal{N}_\infty^-(\beta) .$$

The above discussion can be summarized in the following lemma.

Lemma 4.2. *Let* $(x^k, s^k) \in \mathcal{N}_\infty^-(\beta)$ *be the k-th iterate of the Algorithm 3.5. If*

$$\| h(\theta) \|_\infty \le (1-\beta)\gamma\theta\mu_k , \tag{4.6}$$

then

$$(x(\theta), s(\theta)) \in \mathcal{N}_\infty^-(\beta) .$$

In order to find a lower bound for stepsize θ_k we need to derive another upper bound for $\| h(\theta) \|_\infty$ different from the one given in (4.6). We use the modified scaled Lipschitz condition (SLC).

Lemma 4.3. *If*

$$\theta \| (X^k)^{-1} \Delta x \|_\infty \le \alpha < 1 ,$$

then

$$\| h(\theta) \|_\infty \le \bar{v}_\psi(\alpha)\theta^2 \| (D^k)^{-1} \Delta x \| \| D^k \Delta s \| , \tag{4.7}$$

where

$$\bar{v}_\psi(\alpha) = (1 + (1+\alpha) v_\psi(\alpha)), \quad D^k = (X^k)^{1/2} (S^k)^{-1/2} . \tag{4.8}$$

Proof: Recall the definition (3.14) of $h(\theta)$

$$h(\theta) = \theta^2 \Delta X \Delta s + (X^k + \theta\Delta X)(\psi(x(\theta)) - \psi(x^k) - \theta\nabla\psi(x^k)\Delta x) .$$

Since ψ satisfied SLC, and $\theta \in (0,1)$ we conclude that if

$$\theta \| (X^k)^{-1} \Delta x \|_\infty \le \alpha < 1 ,$$

then

$$| x_i^k + \theta\Delta x_i | \le (1+\alpha) x_i^k < 2 x_i^k, \quad i = 1, \ldots, n+1 ,$$

and using also (4.2)

$$\begin{aligned}
\| h(\theta) \|_\infty &\le \theta^2 \| \Delta X \Delta s \|_\infty + \| (X^k + \theta\Delta X)(\psi(x(\theta)) - \psi(x^k) - \theta\nabla\psi(x^k)\Delta x) \|_\infty \\
&\le \theta^2 \| \Delta X \Delta s \|_\infty + (1+\alpha) \| X^k (\psi(x(\theta)) - \psi(x^k) - \theta\nabla\psi(x^k)\Delta x) \|_\infty \\
&\le \theta^2 \| \Delta X \Delta s \|_\infty + (1+\alpha) v_\psi(\alpha)\theta^2 | \Delta x^T \nabla\psi(x^k) \Delta s | \\
&\le \theta^2 \| \Delta X \Delta s \|_\infty + (1+\alpha) v_\psi(\alpha)\theta^2 | (\Delta x)^T \Delta s | \\
&= \theta^2 \| \Delta X (D^k)^{-1} D^k \Delta s \|_\infty + (1+\alpha) v_\psi(\alpha)\theta^2 | (\Delta x)^T (D^k)^{-1} D^k \Delta s | \\
&\le \theta^2 \| (D^k)^{-1} \Delta x \| \| D^k \Delta s \| + (1+\alpha) v_\psi(\alpha)\theta^2 \| (D^k)^{-1} \Delta x \| \| D^k \Delta s \| \\
&= (1 + (1+\alpha) v_\psi(\alpha))\theta^2 \| (D^k)^{-1} \Delta x \| \| D^k \Delta s \| .
\end{aligned}$$

♦

From the above lemma we conclude that the problem of finding upper bound on $\| h(\theta) \|_\infty$ is reduced to the problem of finding upper bounds on $\| (D^k)^{-1} \Delta x \|$ and $\| D^k \Delta s \|$. In order to do so we need several technical lemmas. The first one is proposition 2.2 of Ji et al. [10] which gives error bounds for a system of the type (3.7)-(3.8).

Lemma 4.4. *Let $x, s, \bar{a}, \bar{b}$ be four vectors of the same dimension with $(x,s) > 0$, and let M be a $P_*(\kappa)$-matrix. The solution (u,v) of the linear system*

$$Su + Xv = \bar{a}\ , \tag{4.9}$$
$$Mu - v = \bar{b}\ , \tag{4.10}$$

satisfies the following inequalities

$$\| D^{-1} u \| \le \| b \| + \sqrt{\| a \|^2 + \| b \|^2 + 2\kappa \| c \|^2}\ , \tag{4.11}$$
$$\| Dv \| \le \sqrt{\| a \|^2 + \| b \|^2 + 2\kappa \| c \|^2}\ , \tag{4.12}$$
$$\| D^{-1} u \|^2 + \| Dv \|^2 \le \| a \|^2 + 2 \| b \|^2 + 2\kappa \| c \|^2 + 2 \| b \| \sqrt{\| a \|^2 + \| b \|^2 + 2\kappa \| c \|^2} = \chi^2\ , \tag{4.13}$$
$$\| Uv \| \le \frac{1}{8} \| a \|^4 + \frac{1}{4} \chi^2 (\chi^2 - \| a \|^2)\ , \tag{4.14}$$

where

$$D = X^{1/2} S^{-1/2},\quad a = (XS)^{-1/2} \bar{a},\quad b = D\bar{b},\quad c = a + b\ . \tag{4.15}$$

In particular, for the system (3.7)-(3.8) we have

$$a = (X^k S^k)^{-1/2} (\gamma \mu_k e - X^k s^k),\quad b = \eta D^k r^k\ . \tag{4.16}$$

Hence the problem of finding upper bounds on $\| (D^k)^{-1} \Delta x \|$ and $\| D^k \Delta s \|$ is further reduced to the problem of finding upper bounds on $\| a \|$ and $\| b \|$ defined above. In order to find them we need to establish the boundedness of the iteration sequence (x^k, s^k) produced by the Algorithm 3.5.

Lemma 4.5. *Let $(x^0, s^0) > 0$ be the initial point and let $(x^k, s^k) > 0$ be the k-th iterate of the Algorithm 3.5. Then*

$$(x^k)^T s^0 + (s^k)^T x^0 \le 2(1 + 4\kappa)(x^0)^T s^0\ . \tag{4.17}$$

Proof: In what follows we denote

$$\Theta_k = \prod_{i=0}^{k-1} (1 - \eta \theta_i)\ . \tag{4.18}$$

Then from (4.3) and (4.4) we have

$$\mu_k = \Theta_k \mu_0,\quad r^k = \Theta_k r^0\ . \tag{4.19}$$

Using (3.6) we obtain

$$\begin{aligned}(x^k)^T s^0 + (s^k)^T x^0 &= (x^k)^T (r^0 + \psi(x^0)) + (x^0)^T (r^k + \psi(x^k)) \\ &= (x^k)^T r^0 + (x^0)^T r^k + (x^k)^T \psi(x^0) + (x^0)^T \psi(x^k).\end{aligned} \tag{4.20}$$

First we estimate the term $(x^k)^T r^0 + (x^0)^T r^k$ in (4.20). From (4.19), (3.6) and (3.2) we have

$$\Theta_k (x^k)^T r^0 = (x^k)^T r^k = (x^k)^T (s^k - \psi(x^k)) = (x^k)^T s^k . \tag{4.21}$$

So

$$(x^k)^T r^0 + (x^0)^T r^k = (1 + \Theta_k)(x^0)^T s^0 . \tag{4.22}$$

Next we need to estimate the second term in (4.20), i.e. $(x^k)^T \psi(x^0) + (x^0)^T \psi(x^k)$. Using (3.2) and the fact that ψ is a homogeneous function of order 1 we conclude

$$\begin{aligned}&-\Theta_k((x^k)^T \psi(x^0) + (x^0)^T \psi(x^k)) \\ &= (x^k)^T \psi(x^k) + \Theta_k (x^0)^T \psi(\Theta_k x^0) - \Theta_k((x^k)^T \psi(x^0) + (x^0)^T \psi(x^k)) \\ &= (x^k - \Theta_k x^0)^T (\psi(x^k) - \psi(\Theta_k x^0)).\end{aligned} \tag{4.23}$$

On the other hand, from (3.6) and (4.19) we have

$$\psi(x^k) - \psi(\Theta_k x^0) = (s^k - r^k) - \Theta_k (s^0 - r^0) = s^k - \Theta_k s^0 . \tag{4.24}$$

Using (4.19), (4.24), positivity of (x^k, s^k) and the fact that ψ is a P_*-function, we obtain

$$\begin{aligned}(x^k - \Theta_k x^0)^T (\psi(x^k) - \psi(\Theta_k x^0)) &\geq -4\kappa \sum_{i \in \mathfrak{T}_\psi^+} (x_i^k - \Theta_k x_i^0)(\psi_i(x^k) - \psi_i(\Theta_k x^0)) \\ &= -4\kappa \sum_{i \in \mathfrak{T}_\psi^+} (x_i^k - \Theta_k x_i^0)(s_i^k - \Theta_k s_i^0) \\ &= -4\kappa \sum_{i \in \mathfrak{T}_\psi^+} (x_i^k s_i^k + \Theta_k^2 x_i^0 s_i^0 - \Theta_k (x_i^0 s_i^k + x_i^k s_i^0)) \\ &\geq -4\kappa \sum_{i \in \mathfrak{T}_\psi^+} (x_i^k s_i^k + \Theta_k^2 x_i^0 s_i^0) \\ &\geq -4\kappa ((x^k)^T s^k + \Theta_k^2 (x^0)^T s^0) \\ &= -4\kappa \Theta_k (1 + \Theta_k)(x^k)^T s^k.\end{aligned} \tag{4.25}$$

From (4.23) and (4.25) we derive

$$(x^0)^T \psi(x^k) + (x^k)^T \psi(x^0) \geq 4\kappa (1 + \Theta_k)(x^k)^T s^k . \tag{4.26}$$

Substituting (4.22) and (4.26) into (4.20) we obtain

$$(x^k)^T s^0 + (s^k)^T x^0 \le (1+4\kappa)(1+\Theta_k)(x^k)^T s^k \le 2(1+4\kappa)(x^0)^T s^0.$$

The last inequality above is due to the fact that $\Theta_k \in (0,1)$. ♦

Now we are able to obtain upper bounds for $\| a \|$ and $\| b \|$ defined by (4.16).

Lemma 4.6. *Let* $(x^k, s^k) > 0$ *be the k-th iterate of the Algorithm 3.5. We set the constants in the algorithm as follows*

$$\gamma \le 2\beta, \quad \eta = \frac{\rho}{\sqrt{n+1}}, \quad 0 < \rho < \sqrt{n+1} . \tag{4.27}$$

Then for $\| a \|$ *and* $\| b \|$ *defined by (4.16) we have*

$$\| a \| \le \sqrt{(n+1)\mu_k} , \tag{4.28}$$

$$\| b \| \le \delta(1+4\kappa)\sqrt{(n+1)\mu_k} , \tag{4.29}$$

where

$$\delta = 2\frac{\rho}{\sqrt{\beta}} \| (S^0)^{-1} r^0 \|_\infty . \tag{4.30}$$

Proof: The proof of the (4.28) is the same as the proof of Lemma 3.4 in [31]. The proof of (4.29) is as follows. We have

$$\| b \| = \| \eta D^k r^k \| = \eta \| (X^k S^k)^{-1/2} X^k r^k \| \le \eta \| (X^k S^k)^{-1/2} \| \| X^k r^k \| .$$

Using (4.19) and the fact that $(x^k, s^k) \in \mathcal{N}_\infty^-(\beta)$ we obtain

$$\begin{aligned} \| b \| &\le \frac{\eta \Theta_k}{\sqrt{\beta \mu_k}} \| X^k r^0 \| \\ &\le \frac{\eta \Theta_k}{\sqrt{\beta \mu_k}} \| X^k r^0 \|_1 \\ &\le \frac{\eta \Theta_k}{\sqrt{\beta \mu_k}} \| (S^0)^{-1} r^0 \|_\infty \| X^k s^0 \|_1 \\ &= \frac{\eta \Theta_k}{\sqrt{\beta \mu_k}} \| (S^0)^{-1} r^0 \|_\infty (x^k)^T s^0. \end{aligned}$$

By virtue of Lemma 4.5 we obtain

$$\begin{aligned} \| b \| &\le \frac{\eta \Theta_k}{\sqrt{\beta \mu_k}} \| (S^0)^{-1} r^0 \|_\infty \, 2(1+4\kappa)(x^0)^T s^0 \\ &= 2\frac{\eta}{\sqrt{\beta}} \| (S^0)^{-1} r^0 \|_\infty \, (1+4\kappa)(n+1)\sqrt{\mu_k} \\ &= 2\frac{\rho}{\sqrt{\beta}} \| (S^0)^{-1} r^0 \|_\infty \, (1+4\kappa)\sqrt{(n+1)\mu_k}. \end{aligned}$$

♦

Observe that, since $\gamma = 1-\eta$, the requirements (4.30) in the above lemma lead to the conclusion

$$1-2\beta \le \rho < \sqrt{n+1}\,. \tag{4.31}$$

From the above lemma and Lemma 4.4 the upper bounds for $\|(D^k)^{-1}\Delta x\|$ and $\|D^k \Delta s\|$ follow easily.

Corollary 4.7. *Let* $(\Delta x, \Delta s)$ *be the direction calculated in Algorithm 3.5 and let the constants in the algorithm be chosen as in (4.27). Then*

$$\|(D^k)^{-1}\Delta x\| \le \delta_1 (1+4\kappa)^{3/2}\sqrt{(n+1)\mu_k}\,, \tag{4.32}$$

$$\|D^k \Delta s\| \le \delta_2 (1+4\kappa)^{3/2}\sqrt{(n+1)\mu_k}\,, \tag{4.33}$$

where

$$\delta_1 = \delta + \sqrt{1+\delta^2}, \quad \delta_2 = \sqrt{1+\delta^2}\,. \tag{4.34}$$

Proof: We have

$$\begin{aligned}
\|(D^k)^{-1}\Delta x\| &\le \|b\| + \sqrt{\|a\|^2 + \|b\|^2 + 2\kappa\|c\|^2} \\
&\le \|b\| + \sqrt{1+2\kappa}\sqrt{\|a\|^2 + \|b\|^2} \\
&\le (\delta(1+4\kappa) + \sqrt{1+2\kappa}\sqrt{1+(\delta(1+4\kappa))^2})\sqrt{(n+1)\mu_k} \\
&\le (\delta + \sqrt{1+\delta^2})(1+4\kappa)^{3/2}\sqrt{(n+1)\mu_k}.
\end{aligned}$$

Similarly

$$\begin{aligned}
\|D^k \Delta s\| &\le \sqrt{\|a\|^2 + \|b\|^2 + 2\kappa\|c\|^2} \\
&\le \sqrt{1+2\kappa}\sqrt{\|a\|^2 + \|b\|^2} \\
&\le \sqrt{1+2\kappa}\sqrt{1+(\delta(1+4\kappa))^2}\sqrt{(n+1)\mu_k} \\
&\le \sqrt{1+\delta^2}(1+4\kappa)^{3/2}\sqrt{(n+1)\mu_k}.
\end{aligned}$$

♦

Now we have all the ingredients to prove linear global convergence of the iteration sequence produced by Algorithm 3.5.

Theorem 4.8. *Algorithm 3.5 with the following choice of parameters*

$$2\sqrt{\beta} \le \alpha < 1, \quad 1-2\beta \le \rho < \sqrt{n+1}, \quad \eta = \frac{\rho}{\sqrt{n+1}}, \quad \gamma = 1-n\,, \tag{4.35}$$

finds ε *-approximate solution of HNCP in at most*

$$O\left((n+1)\overline{\upsilon}_\psi(\alpha)\overline{\delta}(1+4\kappa)^3 \max\left\{ \ln\frac{(x^0)^T s^0}{\varepsilon}, \ln\frac{|r^0|}{\varepsilon} \right\} \right)$$

iterations, where $\overline{\upsilon}_\psi(\alpha)$ *is defined by (4.8), and* $\overline{\delta} = \delta_1\delta_2$, *where* δ_1, δ_2 *are defined by (4.34).*

Proof: Substituting (4.32) and (4.33) into (4.7) we obtain

$$\| h(\theta) \|_\infty \le \overline{\upsilon}_\psi(\alpha)\theta^2 \overline{\delta}(1+4\kappa)^3(n+1)\mu_k \,, \tag{4.36}$$

where $\overline{\delta} = \delta_1\delta_2$. Comparing (4.6) and (4.36) we derive

$$\hat{\theta} = \frac{(1-\beta)\gamma}{\overline{\upsilon}_\psi(\alpha)\overline{\delta}(1+4\kappa)^3(n+1)} \,, \tag{4.37}$$

provided that

$$\hat{\theta} \| (X^k)^{-1}\Delta x \|_\infty \le \alpha \tag{4.38}$$

holds. To assure (4.38) we use (4.32) and the fact that $(x^k, s^k) \in \mathcal{N}_\infty^-(\beta)$,

$$\| (X^k)^{-1}\Delta x \|_\infty \le \| (X^k)^{-1}\Delta x \| \tag{4.39}$$

$$= \| (X^k S^k)^{-1/2} (D^k)^{-1} \Delta x \| \tag{4.40}$$

$$\le \frac{1}{\sqrt{\beta\mu_k}} \| (D^k)^{-1}\Delta x \| \tag{4.41}$$

$$\le \frac{\delta_1}{\sqrt{\beta}}(1+4\kappa)^3\sqrt{n+1} \,. \tag{4.42}$$

Substituting (4.42) into (4.38) we obtain

$$\frac{(1-\beta)\gamma}{\overline{\upsilon}_\psi \delta_2 (1+4\kappa)^3 \sqrt{(n+1)\beta}} \le \alpha \,.$$

Since $\overline{\upsilon}_\psi > 1$, $\delta_1 > 1$, $\delta_2 > 1$, $\beta < 1$ and $\gamma \le 2\beta$, the above inequality implies

$$2\sqrt{\beta} \le \alpha < 1 \,. \tag{4.43}$$

Choosing parameters as in (4.43) will guarantee that (4.38) holds, and therefore by Lemma 4.3 the inequality (4.36) will hold for θ defined in (4.37), i.e.

$$\| h(\hat{\theta}) \|_\infty \le \overline{\upsilon}_\psi(\alpha)\hat{\theta}^2 \overline{\delta}(1+4\kappa)^3(n+1)\mu_k = (1-\beta)\gamma\hat{\theta}\mu_k \,.$$

From Lemma 4.2 then follows

$$(x(\hat{\theta}), s(\hat{\theta})) \in \mathcal{N}_\infty^-(\beta) \,.$$

The selection of the stepsize, as described in the Algorithm 3.5, together with (4.3), implies

$$\mu_{k+1} = \mu(\theta_k) \le \mu(\hat{\theta}) = (1-\eta\hat{\theta})\mu_k \le (1-\eta\hat{\theta})^{k+1}\mu_0 ,$$

and similarly

$$r^{k+1} \le (1-\eta\hat{\theta})^{k+1} r^0 .$$

Hence Algorithm 3.5 requires

$$O\left((n+1)\bar{\upsilon}_\psi(\alpha)\bar{\delta}(1+4\kappa)^3 \max\left\{ \ln\frac{(x^0)^T s^0}{\varepsilon}, \ln\frac{|r^0|}{\varepsilon} \right\} \right)$$

iterations to obtain ε -approximate solution for HNCP.

5. LOCAL CONVERGENCE

In this section we prove that a sequence $\{\mu_k\}$, generated by a modified Algorithm 3.5, converge to zero with R-order and Q-order at least 2, while sequence $\{(x^k, s^k)\}$ converges to a strictly complementary solution (we made assumption that it exists) with R-order at least 2. Below we recall the definitions of Q-order and R-order convergence. For more details see Potra [27].

A positive sequence $\{a_k\}$ is said to converge to zero with Q-order at least $t > 1$ if there exists a constant $c \ge 0$ such that

$$a_{k+1} \le c a_k^t, \quad \forall k . \tag{5.1}$$

The above sequence converges to zero with Q-order exactly t if

$$t = \sup\{\bar{t} : \{a_k\} \text{ converges with Q-order at least } \bar{t}\} , \tag{5.2}$$

or equivalently iff

$$t = \liminf_{k\to\infty} \frac{\log a_{k+1}}{\log a_k} . \tag{5.3}$$

A positive sequence $\{a_k\}$ is said to converge to zero with R-order at least $t > 1$ if there exists a constant $c \ge 0$ and a constant $b \in (0,1)$ such that

$$a_{k+1} \le c b^{t^k}, \quad \forall k . \tag{5.4}$$

The key part in proving the local convergence result is relating the components of the iteration sequence (x^k, s^k) generated by Algorithm 3.5 to the primal-dual gap $(x^k)^T s^k$. We have the following lemma.

Lemma 5.1. *Let (x^*, s^*) be a strictly complementary solution of HNCP, and let (x^k, s^k) be the k-th iterate of the Algorithm 3.5. Then*

$$(x^*)^T s^k + (s^*)^T x^k \le \varphi (x^k)^T s^k , \tag{5.5}$$

where φ is defined by (5.6).

Proof: Using (4.18), (4.19), (4.21), positivity of the initial point $(x^0, s^0) > 0$, P_*-property of ψ, and the fact that $s^* = \psi(x^*)$, $(x^*)^T s^* = 0$, $(x^*, s^*) \ge 0$ we derive

$$\begin{aligned}
&(x^*)^T s^k + (s^*)^T x^k = \\
&= -(x^k - x^*)^T (s^k - s^*) + (x^k)^T s^k \\
&= -(x^k - x^*)^T (\psi(x^k) - \psi(x^*)) - (r^k)^T (x^k - x^*) + (x^k)^T s^k \\
&\le 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} (x_i^k - x_i^*)(\psi_i(x^k) - \psi_i(x^*)) - (r^k)^T x^k + (r^k)^T x^* + (x^k)^T s^k \\
&= 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} (x_i^k - x_i^*)(s_i^k - r_i^k - s_i^*) - (x^k)^T s^k + \Theta_k (r^0)^T x^* + (x^k)^T s^k \\
&= 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} (x_i^k - x_i^*)(s_i^k - s_i^*) - 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} r_i^k (x_i^k - x_i^*) + \Theta_k \sum_{i=1}^{n+1} r_i^0 x_i^* \\
&= 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} (x_i^k s_i^k - x_i^k s_i^* - x_i^* s_i^k + x_i^* s_i^*) - 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} r_i^k (x_i^k - x_i^*) + \Theta_k \sum_{i=1}^{n+1} r_i^0 x_i^0 \frac{x_i^*}{x_i^0} \\
&\le 4\kappa (x^k)^T s^k - 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} r_i^k (x_i^k - x_i^*) + \| (X^0)^{-1} x^* \|_\infty \Theta_k \sum_{i=1}^{n+1} r_i^0 x_i^0 \\
&= 4\kappa (x^k)^T s^k + 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} r_i^k (x_i^* - x_i^k) + \| (X^0)^{-1} x^* \|_\infty \Theta_k (r^0)^T x^0 \\
&= 4\kappa (x^k)^T s^k + 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} r_i^k (x_i^* - x_i^k) + \| (X^0)^{-1} x^* \|_\infty \Theta_k (x^0)^T s^0 \\
&\le 4\kappa (x^k)^T s^k + 4\kappa \sum_{i \in \mathfrak{T}_\psi^+} \Theta_k | r_i^0 | | x_i^* - x_i^k | + \| (X^0)^{-1} x^* \|_\infty (x^k)^T s^k \\
&\le 4\kappa (x^k)^T s^k + 4\kappa \Theta_k \sum_{i \in \mathfrak{T}_\psi^+} | r_i^0 x_i^* | + 4\kappa \Theta_k \sum_{i \in \mathfrak{T}_\psi^+} | r_i^0 x_i^k | + \| (X^0)^{-1} x^* \|_\infty (x^k)^T s^k \\
&= 4\kappa (x^k)^T s^k + 4\kappa \Theta_k \| X^* r^0 \|_1 + 4\kappa \Theta_k \| X^k r^0 \|_1 + \| (X^0)^{-1} x^* \|_\infty (x^k)^T s^k \\
&\le 4\kappa (x^k)^T s^k + 4\kappa \Theta_k \| X^* r^0 \|_\infty \frac{n+1}{(x^0)^T s^0} (x^0)^T s^0 + 4\kappa \Theta_k \| (S^0)^{-1} r^0 \|_\infty \| X^k s^0 \|_1 \\
&+ \| (X^0)^{-1} x^* \|_\infty (x^k)^T s^k \\
&= 4\kappa (x^k)^T s^k + 4\kappa \frac{1}{\mu_0} \| X^* r^0 \|_\infty (x^k)^T s^k + 4\kappa \Theta_k \| (S^0)^{-1} r^0 \|_\infty (x^k)^T s^0 \\
&+ \| (X^0)^{-1} x^* \|_\infty (x^k)^T s^k
\end{aligned}$$

$$\le 4\kappa(x^k)^T s^k + 4\kappa \frac{1}{\mu_0} \| X^* r^0 \|_\infty (x^k)^T s^k + 4\kappa\Theta_k \| (S^0)^{-1} r^0 \|_\infty 2(1+4\kappa)(x^0)^T s^0$$
$$+ \| (X^0)^{-1} x^* \|_\infty (x^k)^T s^k$$
$$= \left(\| (X^0)^{-1} x^* \|_\infty + 4\kappa \left(1 + \frac{1}{\mu_0} \| X^* r^0 \|_\infty + 2(1+4\kappa) \| (S^0)^{-1} r^0 \|_\infty \right) \right) (x^k)^T s^k.$$

If we denote

$$\begin{aligned} \zeta &= \| (X^0)^{-1} x^* \|_\infty, \\ \rho &= \frac{1}{\mu_0} \| X^* r^0 \|_\infty, \\ \nu &= \| (S^0)^{-1} r^0 \|_\infty, \\ \varphi &= \zeta + 4\kappa(1 + \rho + 2(1+4\kappa)\nu), \end{aligned} \tag{5.6}$$

then we obtain (5.5). ♦

It has been shown that for LP a unique partition $\{B, N\}$ of the set $\{1,\ldots,n\}$ exists such that

(i) there exists a solution (x^*, s^*) with

$$x_B^* > 0, \quad s_N^* > 0, \tag{5.7}$$

(ii) for each solution (x, s)

$$x_N = 0, \quad s_B = 0. \tag{5.8}$$

The result has been generalized for LCP with the assumption that a strict complementarity solution exists (even for P_* case). Potra and Ye [30] showed that the same is true for NCP.

Suppose that NCP has a strictly complementary solution and let $\{B_f, N_f\}$ be the above mentioned partition. Then by virtue of Lemma 3.2 (i)

$$B = B_f \cup \{\text{index for } \tau^*\}, \tag{5.9}$$
$$N = N_f \cup \{\text{index for } \sigma^*\} \tag{5.10}$$

is a partition for HNCP. Now we are ready to prove the following important lemma

Lemma 5.2. *Suppose that HNCP has a strictly complementary solution* (x^*, s^*). *Let* (x^k, s^k) *be the k-th iterate of the Algorithm 3.5. There exist three positive constants*

$$\xi = \frac{(n+1)\varphi}{\min\{\min_{i\in B} x_i^*, \min_{i\in N} s_i^*\}}, \tag{5.11}$$

$$\phi = \frac{\beta}{\xi}, \tag{5.12}$$

$$\vartheta = \frac{2(1+4\kappa)(x^0)^T s^0}{\min\{\min_{i\in B} x_i^0, \min_{i\in N} s_i^0\}}, \tag{5.13}$$

such that

$$\phi \le x_i^k \le \vartheta \quad s_i^k \le \xi\mu_k, \quad \forall i \in B, \tag{5.14}$$

$$\phi \le s_i^k \le \vartheta \quad x_i^k \le \xi\mu_k, \quad \forall i \in N. \tag{5.15}$$

Proof: Using Lemma 5.1 and partition $\{B, N\}$ we obtain

$$\sum_{i\in B} x_i^* s_i^k + \sum_{i\in N} s_i^* x_i^k \le \varphi (x^k)^T s^k .$$

Since $(x^k, s^k) \in \mathcal{N}_\infty^-(\beta)$, from the above inequality we deduce for each $i \in B$

$$x_i^k = \frac{x_i^k s_i^k}{s_i^k} \ge \frac{\beta\mu_k}{s_i^k} = \frac{\beta}{n+1}\frac{(x^k)^T s^k}{s_i^k} \ge \frac{\beta}{\xi} = \phi ,$$

and

$$s_i^k \le \frac{\varphi (x^k)^T s^k}{x_i^*} \le \xi\mu_k .$$

Also, an immediate consequence of Lemma 4.5 is

$$x_i^k \le \frac{2(1+4\kappa)(x^0)^T s^0}{s_i^0} \le \vartheta, \quad \forall i \in \{1,\ldots,n+1\} .$$

Thus (5.14) is proved. Similarly we prove (5.15).

An immediate consequence of the above lemma is the following corollary:

Corollary 5.3. *Any accumulation point* $(x^* s^*)$ *of the sequence obtained by Algorithm 3.5 is a strictly complementary solution of HNCP.*

The above corollary together with (5.9), (5.10) assures that a strictly complementary solution of HNCP will be of the type as in Lemma 3.2 (ii), thus enabling us to find a strictly complementary solution of NCP.

To prove the local convergence result we modify Algorithm 3.5 in such a way that for a sufficiently large k, say K, we set $\gamma = 0$, i.e. centering part of the direction is omitted and only an affine-scaling direction is calculated. Hence the algorithm becomes an affine-scaling algorithm or, in other words, a damped Newton method. The existence of the treshold value K will be established later in the text. For now, without the loss of generality, we can assume $K = 0$.

In addition, instead of keeping a fixed neighborhood of the central path we enlarge it at each iteration. Let

$$\beta_0 = \beta, \quad \beta_{k+1} = \beta_k - \pi_k, \quad \forall k , \tag{5.16}$$

where

$$\sum_{k=0}^{\infty} \pi_k < \infty, \quad \pi_k > 0, \quad \forall k . \tag{5.17}$$

A particular choice of π_k is as in [31]

$$\pi_k = \frac{\beta}{3^{k+1}} . \tag{5.18}$$

Thus

$$\frac{\beta}{2} < \cdots < \beta_{k+1} < \beta_k < \cdots < \beta_0 = \beta ,$$

and

$$\mathcal{N}_\infty^-(\beta) \subseteq \cdots \mathcal{N}_\infty^-(\beta_k) \subseteq \mathcal{N}_\infty^-(\beta_{k+1}) \subseteq \cdots \mathcal{N}_\infty^-(\beta/2) . \tag{5.19}$$

With the above modifications Algorithm 3.5 is reduced to the following affine-scaling algorithm.

Algorithm 5.4

I (*Initialization*)

Let $\varepsilon > 0$ *be a given tolerance, and let* $\beta \in (0,1)$ *be the given constant. Set* $\beta_0 = \beta$ *. Suppose starting point* $(x^0, s^0) \in \mathcal{N}_\infty^-(\beta_0)$ *is available. Calculate* $\mu_0 = (x^0)^T s^0 / (n+1)$ *and set* $k = 0$ *.*

S (*Step*)

Given $(x^k, s^k) \in \mathcal{N}_\infty^-(\beta_k)$ *solve the system*

$$\nabla \psi(x^k) \Delta x - \Delta s = r^k , \tag{5.20}$$

$$S^k \Delta x + X^k \Delta s = -X^k s^k . \tag{5.21}$$

Let

$$x(\theta) = x^k + \theta \Delta x, \quad s(\theta) = \psi(x(\theta)) + (1-\theta) r^k , \tag{5.22}$$

and perform a line search to determine the maximal stepsize $0 < \theta_k < 1$ *such that*

$$(x(\theta_k), s(\theta_k)) \in \mathcal{N}_\infty^-(\beta_{k+1}) , \tag{5.23}$$

and $\mu(\theta_k)$ *minimizes* $\mu(\theta)$. *Set*

$$x^{k+1} = x(\theta_k), \quad s^{k+1} = s(\theta_k) , \tag{5.24}$$

and

$$\beta_{k+1} = \beta_k - \frac{\beta}{3^{k+1}} . \tag{5.25}$$

T (*Termination*)

If

$$(x^{k+1}, s^{k+1}) \in \Psi_\varepsilon = \{(x,s) \geq 0 : x^T s \leq \varepsilon, \| s - \psi(x) \| \leq \varepsilon\}, \tag{5.26}$$

then stop, otherwise set $k := k+1$ *and go to (S).*

A similar modification was employed in [37] on the predictor-corrector algorithm for the monotone LCP, in [30] on the potential reduction algorithm for monotone NCP and in [31] on the path following algorithm for monotone NCP. In the linear case, i.e. for LCP, the above modifications, together with the existence of a strict complementary solution, were necessary and sufficient to prove the local convergence. In the nonlinear case certain additional assumption on the nonsingularity of the Jacobian submatrix is necessary. We adopt Assumption 2 from [31].

Nonsingularity of the Jacobian submatrix (NJS)

Let the Jacobian matrix $\nabla\psi$ be partitioned as follows

$$\nabla\psi(x) = \begin{bmatrix} \nabla\psi_{BB}(x) & \nabla\psi_{BN}(x) \\ \nabla\psi_{NB}(x) & \nabla\psi_{NN}(x) \end{bmatrix}, \tag{5.27}$$

where $\{B, N\}$ is partition of HNCP described by (5.7)-(5.10). We assume that matrix $\nabla\psi_{BB}$ is nonsingular on the following compact set

$$\Gamma = \{x \geq 0 : x_B \geq \phi e_B, x \leq \vartheta e\}, \tag{5.28}$$

where ϕ and ϑ are defined in Lemma 5.2. ♦

So far we have made the following assumptions:

- function ψ is a P_*-function,
- function ψ satisfies the scaled Lipsschitz condition (SLC),
- the existence of the a strict complementary solution (ESCS),
- nonsingularity of the Jacobian submatrix (NJS),

and we assume they hold throughout this section.

Since in this section $\gamma = 0$, i.e. $\eta = 1$, equations (4.3) and (4.4) are reduced to

$$\mu_{k+1} = (1-\theta_k)\mu_k, \quad r^{k+1} = (1-\theta_k)r^k. \tag{5.29}$$

If we are able to prove

$$1 - \theta_k = O(\mu_k),$$

the local convergence result would follow. In order to do so we need to revisit the analysis performed for the global convergence and adjust it according to the modification and assumptions made above.

Note first that the lemmas proved so far in this section remain valid for Algorithm 5.4. Next we show that the direction calculated in the algorithm is bounded from above by μ_k.

Lemma 5.5. *Let* $(\Delta x, \Delta s)$ *be a solution of the system (5.20)-(5.21). Then*

$$\| \Delta x \| \le c_0 \mu_k, \quad \| \Delta s \| \le c_0 \mu_k , \tag{5.30}$$

where c_0 *is a constant independent of* k.

Proof: First we show that

$$\| (\Delta x)_N \| \le c_0' \mu_k, \quad \| (\Delta s_B) \| \le c_0' \mu_k, \tag{5.31}$$

for some constant c_0' independent of k. We have

$$\begin{aligned} \| (\Delta x)_N \| &= \| D_N^k (D_N^k)^{-1} (\Delta x)_N \| \\ &\le \| D_N^k \| \| (D_N^k)^{-1} (\Delta x)_N \| \\ &\le \| D_N^k \| \delta_1 (1+4\kappa)^{3/2} \sqrt{n+1} \sqrt{\mu_k}. \end{aligned}$$

The last inequality above is due to (4.32). Next, we need to estimate $\| D_N^k \|$. Using (5.15) we obtain

$$\| D_N^k \| = \max_{i \in N} \sqrt{\frac{x_i^k}{s_i^k}} \le \sqrt{\frac{\xi \mu_k}{\phi}} .$$

Hencey

$$\| (\Delta x)_N \| \le \left(\sqrt{\frac{\xi}{\phi}} \delta_1 (1+4\kappa)^{3/2} \sqrt{n+1} \right) \mu_k .$$

Similarly, by virtue of (4.33) and (5.14) we have

$$\| (\Delta s)_B \| \le \left(\sqrt{\frac{\xi}{\phi}} \delta_2 (1+4\kappa)^{3/2} \sqrt{n+1} \right) \mu_k .$$

Since by (4.34) $\delta_1 \ge \delta_2$ we can set

$$c_0' = \sqrt{\frac{\xi}{\phi}} \delta_1 (1+4\kappa)^{3/2} \sqrt{n+1} ,$$

and (5.31) is proved.

We still need to prove

$$\| (\Delta x)_B \| \le \bar{c}_0 \mu_k, \quad \| (\Delta s)_N \| \le \bar{c}_0 \mu_k , \tag{5.32}$$

for some constant $\bar{c}_0$ independent of k. Using (5.27) equation (5.20) can be partitioned into system

$$\nabla \psi_{BB}(x^k)(\Delta x)_B + \nabla \psi_{BN}(x^k)(\Delta x)_N - (\Delta s)_B = -r_B^k, \tag{5.33}$$

$$\nabla \psi_{NB}(x^k)(\Delta x)_B + \nabla \psi_{NN}(x^k)(\Delta x)_N - (\Delta s)_N = -r_N^k, \tag{5.34}$$

Hence

$$\|(\Delta x)_B\| \le \|(\nabla \psi_{BB}(x^k))^{-1}\| (\|(\Delta s)_B\| + \|\nabla \psi_{BN}(x^k)\| \|(\Delta x)_N\| + \|r_B^k\|), \tag{5.35}$$

$$\|(\Delta s)_N\| \le \|\nabla \psi_{NB}(x^k)\| \|(\Delta x)_B\| + \|\nabla \psi_{NN}(x^k)\| \|(\Delta x)_N\| + \|r_N^k\|. \tag{5.36}$$

From Lemma 5.2 it follows that the iterates (x^k, s^k) of the Algorithm 5.4 belong to the set Γ defined by (5.28). By (NJS) assumption $\nabla \psi_{BB}(x^k)$ is nonsingular on Γ. Thus, since Γ is compact, all matrices above are uniformly bounded. Also from (5.29) we have

$$\frac{1}{\mu_k} r^k = \frac{1}{\mu_0} r^0$$

or

$$\|r^k\| = \frac{\|r^0\|}{\mu_0} \mu_k. \tag{5.37}$$

Using the uniform boundedness of the matrices and substituting (5.31) and (5.37) into (5.35) and (5.36) we obtain (5.32) completing the proof of the lemma. ♦

Lemma 5.6. *There exists a constant* c_1*, independent of* k*, such that*

$$\|h(\theta)\|_\infty \le c_1 \mu_k^2, \tag{5.38}$$

where $h(\theta)$ *is defined by (3.14).*

Proof: From the definition (3.14) of $h(\theta)$ we obtain for each $i \in \{1, \ldots, n+1\}$ the following inequality

$$|h_i(\theta)| \le \theta^2 |(\Delta x)_i (\Delta s)_i| + |x_i^k + \theta(\Delta x)_i| |\psi_i(x^k + \theta \Delta x) - \psi_i(x^k) - \theta \nabla \psi_i(x^k) \Delta x|. \tag{5.39}$$

Recall that ψ satisfies the scaled Lipschitz condition (SLC), i.e. if

$$\theta \|(X^k)^{-1} \Delta x\|_\infty \le \alpha, \tag{5.40}$$

then

$$\|X^k(\psi(x^k + \theta \Delta x) - \psi(x^k) - \theta \nabla \psi(x^k) \Delta x)\|_\infty \le \upsilon(\alpha) \theta^2 |\Delta x^T \nabla \psi(x^k) \Delta x|. \tag{5.41}$$

Substituting (5.40) and (5.41) into (5.39) we obtain

$$|h_i(\theta)| \le \theta^2 \|\Delta x\| \|\Delta s\| + (1+\alpha) \upsilon(\alpha) \theta^2 |\Delta x^T \nabla \psi(x^k) \Delta x|.$$

From Lemma 5.2 it follows that $(x^k, s^k) \in \Gamma$, where Γ is a compact set defined by (5.28). Therefore $\nabla \psi$ is uniformly bounded on Γ, i.e. there exists a constant M such that

$$|h_i(\theta)| \le \theta^2 \| \Delta x \| \| \Delta s \| + (1+\alpha)\upsilon(\alpha)\theta^2 M \| \Delta x \|^2 . \tag{5.42}$$

Using (5.30) and the fact that $\theta < 1$, from (5.42) we derive

$$|h_i(\theta)| \le c_0^2(1 + M(1+\alpha)\upsilon(\alpha))\mu_k^2 , \tag{5.43}$$

providing that (5.40) holds.

To ensure (5.40) we take an index k sufficiently large, i.e. k is the first index, say K_1, such that

$$c_0\mu_{K_1} \le \alpha\phi , \tag{5.44}$$

where ϕ is defined in Lemma 5.2. Using (5.44), (5.30), the fact that $\theta < 1$, and Lemma 5.2 we have for $k \ge K_1$

$$\theta |(\Delta x)_i| \le |(\Delta x)_i| \le \| \Delta x \| \le c_0\mu_k \le \alpha\phi \le \alpha x_i^k, \quad \forall i \in B .$$

Thus, (5.40) holds for sufficiently large k. Hence, we have proved (5.43) but only for $i \in B$.

We still need to prove (5.43) for $i \in N$. Since $\nabla\psi$ is uniformly bounded on Γ, we have

$$\begin{aligned} |\psi_i(x^k + \theta\Delta x) - \psi_i(x^k) - \theta\nabla\psi_i(x^k)\Delta x| &= \left| \int_0^\theta \nabla\psi_i(x^k + t\Delta x)\Delta x\, dt - \theta\nabla\psi_i(x^k)\Delta x \right| \\ &\le 2M\theta \| \Delta x \| \le 2Mc_0\mu_k. \end{aligned} \tag{5.45}$$

Also, using (5.15), (5.30), and the fact that $\theta < 1$, we obtain

$$|x_i^k + \theta(\Delta x)_i| \le (\xi + c_0)\mu_k . \tag{5.46}$$

Substituting (5.45) and (5.46) into (5.39) we get

$$|h_i(\theta)| \le (c_0^2 + 2c_0M(\xi + c_0))\mu_k^2, \quad \forall i \in N . \tag{5.47}$$

From (5.43) and (5.47) we derive (5.38). ♦

Lemma 5.7. *Let (x^k, s^k) be the k-th iterate of the Algorithm 5.4. Define*

$$\hat{\theta}_k = 1 - c_1 \frac{\mu_k}{\pi_k} , \tag{5.48}$$

where c_1 is defined in Lemma 5.6 and π_k is defined by (5.18). Then

$$(x(\hat{\theta}_k), s(\hat{\theta}_k)) \in \mathcal{N}_\infty^-(\beta_{k+1}) . \tag{5.49}$$

Proof: From (4.5), (5.16), (5.29) and (5.38) we obtain

$$\begin{aligned} X(\theta)s(\theta) - \beta_{k+1}\mu(\theta)e &= (1-\theta)X^k s^k + h(\theta) - \beta_{k+1}(1-\theta)\mu_k e \\ &\ge (1-\theta)\beta_k\mu_k e - (\beta_k - \pi_k)(1-\theta)\mu_k e + h(\theta) \\ &\ge \pi_k(1-\theta)\mu_k e - c_1\mu_k^2 e. \end{aligned}$$

If we take θ as in (5.48), then the above inequality implies (5.49).
Note that an immediate consequence of the above lemma is $\hat{\theta}_k \to 1$, which means that Algorithm 5.4 is approaching the pure Newton method.

Now we have all the ingredients to prove the following local convergence result.

Theorem 5.8. *Let $\{(x^k, s^k)\}$ be a sequence generated by the Algorithm 5.4. Then*

(i) $\mu_k \to 0$ with Q-order and R-order at least 2.

(ii) $(x^k, s^k) \to (x^*, s^*)$ with R-order at least 2.

Proof:

(i) Using the rule for selecting stepsize in Algorithm 5.4 and from Lemma 5.7 we have

$$
\begin{aligned}
\mu_{k+1} &= \mu(\theta_k) \\
&\le \mu(\hat{\theta}_k) \\
&= (1-\hat{\theta}_k)\mu_k \\
&= c_1 \frac{\mu_k}{\pi_k} \mu_k \\
&= c_1 \frac{3^{k+1}}{\beta} \mu_k^2 \\
&= \left(\sqrt{c_1 \frac{3^{k+1}}{\beta}} \mu_k \right)^2 \\
&\le \left(\sqrt{3 \frac{c_1}{\beta}} \mu_0 \right)^{2^{k+1}} .
\end{aligned}
\tag{5.50}
$$

Let $k = K_2$ be such that

$$
\sqrt{3 \frac{c_1}{\beta}} \mu_{K_2} < 1 . \tag{5.51}
$$

Now, using (5.44) and (5.51) we can define

$$
K = \max\{K_1, K_2\} , \tag{5.52}
$$

and set

$$
\mu_0 = \mu_K , \quad b = \sqrt{3 \frac{c_1}{\beta}} \mu_0 . \tag{5.53}
$$

Hence, from (5.50) we have

$$
\mu_k \le b^{2^k} , \quad k \ge K . \tag{5.54}
$$

Next, observe that from (5.50) we obtain

$$\log \mu_{k+1} \le 2 \log \mu_k + (k+1)\log 3 + \log(c_1/\beta) , \tag{5.55}$$

and from (5.53) we obtain

$$\log \mu_k \le 2^k \log b < 0 , \tag{5.56}$$

i.e.

$$|\log \mu_k| \ge 2^k |\log b| . \tag{5.57}$$

Thus, using (5.57) we derive

$$\lim_{k\to\infty} \left| \frac{(k+1)\log 3 + \log(c_1/\beta)}{\log \mu_k} \right| \le \lim_{k\to\infty} \frac{|(k+1)\log 3 + \log(c_1/\beta)|}{2^k |\log b|} = 0 . \tag{5.58}$$

Hence, taking into account (5.57) and (5.58), we obtain from (5.55)

$$\liminf_{k\to\infty} \frac{\log \mu_{k+1}}{\log \mu_k} \ge 2 . \tag{5.59}$$

Using definitions (5.3) and (5.4) we conclude from (5.54) and (5.59) that $\mu_k \to 0$ with Q-order and R-order at least 2.

(ii) First we show that $\{(x^k, s^k)\}$ is a Cauchy sequence. Take any $m > k \ge K$. Then

$$\| x^m - x^k \| \le \sum_{i=0}^{m-1} \| x^{k+i+1} - x^{k+i} \| \le \sum_{i=0}^{\infty} \| x^{k+i+1} - x^{k+i} \| . \tag{5.60}$$

Using (5.30) and (5.54) we have

$$\| x^{k+i+1} - x^{k+i} \| = \theta_{k+i} \| \Delta x \| \le c_0 \mu_{k+i} \le c_0 b^{2^{k+i}} . \tag{5.61}$$

Substituting (5.61) into (5.60) we obtain

$$\| x^m - x^k \| \le c_0 b^{2^k} \sum_{i=0}^{\infty} b^{2^i} \le \frac{c_0}{1-b} b^{2^k} . \tag{5.62}$$

We have a similar estimate for s^k proving that $\{(x^k, s^k)\}$ is a Cauchy sequence. Hence the sequence must be convergent, and by (5.9), (5.10) and Corollary 5.3 it converges to such a strict complementary solution (x^*, s^*) of HNCP from which we can derive a strict complementary solution of NCP using Lemma 3.2.

If we let $m \to \infty$, then from (5.62) we obtain

$$\| x^* - x^k \| \le \frac{c_0}{1-b} b^{2^k} , \tag{5.63}$$

and similarly for s^k. Thus, $(x^k, s^k) \to (x^*, s^*)$ with R-order at least 2. ♦

We have proved that if $k \geq K$, where K is the treshold value defined by (5.52), then it is not necessary to calculate the centering part of the direction in the Algorithm 3.5 because the algorithm will produce iterates which are not only centered but also converge to a strictly complementary solution R-quadraticaly. The treshold value K is a theoretical one because some constants used in its calculation may not be known in advance. In practice, as discussed in [37, 31], various heuristic procedures can be developed to determine when to switch from Algorithm 3.5 to Algorithm 5.4. Thus, practical implementation of the algorithm would be a hybrid algorithm which starts with Algorithm 3.5 and then use heuristic "switch time check" procedure to switch to Algorithm 5.4 when suitable.

REFERENCES

[1] Andersen, E., Ye, Y., "On a homogeneous algorithm for the monotone complementarity problem", *Mathematical Programming*, 84 (1999) 375-399.

[2] Anitescu, M., Lesaja, G., and Potra, F.A., "Equivalence between different formulations of the linear complementarity problem", *Optimization Methods and Software*, 7 (1997) 265-290.

[3] Anitescu, M., Lesaja, G., and Potra, F.A., "An infeasible interior-point predictor-corrector algorithm for the p_*-geometric LCP", *Applied Mathematics and Optimization*, 36 (1997) 203-228.

[4] Cottle, R.W., Pang, J.-S., and Stone, R.E., *The Linear Complementarity Problem*, Academic Press, Boston, MA, 1992.

[5] Dikin, I.I., "Iterative solution of problems of linear and quadratic programming", *Soviet Mathematics Doklady*, 8 (1967) 674-675.

[6] Ferris, M.C., and Pang, J.-S., (eds.), *Complementarity and Variational Problems: State of the Art*, SIAM Publishing, Philadelphia, Pennsylvania, 1997.

[7] Güler, O., "Existence of interior points and interior paths in nonlinear monotone complementarity problems", *Mathematics of Operations Research*, 18 (1993) 148-162.

[8] Jansen, B., Roos, K., Terlaky, T., and Yoshise, A., "Polynomiality of primal-dual affine scaling algorithms for nonlinear complementarity problems", *Mathematical Programming*, 78 (1997) 315-345.

[9] Jarre, F., "Interior-point methods via self-concordance or relative Lipschitz condition", *Optimization Methods and Software*, 5 (1995) 75-104.

[10] Ji, J., Potra, F.A., and Sheng, R., "A predictor-corrector method for solving the P_*-matrix LCP from infeasible starting points", *Optimization Methods and Software*, 6 (1995) 109-126.

[11] Kojima, M., Megiddo, N., and Mizuno, S., "A general framework of continuation methods for complementarity problems", *Mathematics of Operations Research*, 18 (4) (1993) 945-963.

[12] Kojima, M., Megiddo, N., Noma, T., and Yoshise, A., "A unified approach to interior point algorithms for linear complementarity problems", *Lecture Notes in Comput. Scie*., 538, 1991.

[13] Kojima, M., Megiddo, N., and Noma, T., "Homotopy continuation method for nonlinear complementarity problems", *Mathematics of Operations Research*, 16 (1991) 754-774.

[14] Kojima, M., Mizuno, S., and Noma, T., "A new continuation method for complementarity problems with uniform P-function", *Mathematical Programming*, 43 (1989) 107-113.

[15] Kojima, M., Mizuno, S., and Noma, T., "Limiting behaviour of trajectories by a continuation method for monotone complementarity problems", *Mathematics of Operations Research*, 15 (1990) 662-675.

[16] Kojima, M., Mizuno, S., and Yoshise, A., "A convex property of monotone complementarity problems", Research Reports on Information Sciences B-267, department of Information

Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152, Japan, March 1993.

[17] Kojima, M., Noma, T., and Yoshise, A., "Global convergence in infeasible-interior-point algorithms", *Mathematical Programming*, 65 (1994) 43-72.

[18] Lesaja, G., "Interior point methods for P_*-complementarity problems", PhD Thesis, University of Iowa, Iowa City, IA, USA, 1996.

[19] McLinden, L., "The analogue of Moreau's proximation theorem, with applications to the nonlinear complementarity problem", *Pacific Journal of Mathematics*, 88 (1980) 101-161.

[20] Miao, J., "A quadratically convergent $O((1+\kappa)\sqrt{n}L)$-iteration algorithm for the $P_*(\kappa)$-matrix linear complementarity problem", Research Report RRR 93, RUTCOR-Rutgers Center for Operations Research, Rutgers University, P.O. Box 5063, New Brunswick, NJ USA, 1993.

[21] Monteiro, R.D.C., Pang, J.-S., and Wang, T., "A positive algorithm for the nonlinear complementarity problem", *SIAM Journal on Optimization*, 5 (1995) 129-148.

[22] Monteiro, R.D.C., and Wright, S.J., "Local convergence of interior-point algorithms for degenerate monotone LCP", *Computational Optimization and Applications*, 3 (1994) 131-155.

[23] Nesterov, Y.E., "Long-step strategies in interior point potential-reduction methods", Working paper, Department SES COMIN, Geneva University, 1993. (to appear in Mathematical Programming).

[24] Nesterov, Y.E., and Nemirovsky, A.S., *Interior Point Polynomial Methods in Convex Programming: Theory and Algorithms*, SIAM Publications, SIAM, Philadelphia, USA, 1994.

[25] Peng, J., Roos, C., and Terlaky, T., "New complexity analysis of primal-dual Newton methods for $P_*(\kappa)$ linear complementarity problems", in: H. Frenk, C. Roos, T. Terlaky, and S. Zhang (eds.), *High Performance Optimization*, Kluwer Academic Publishers, Boston, USA, 1999, 245-265.

[26] Peng, J., Roos, C., Terlaky, T., and Yoshise, A., "Self-regular proximities and new search directions for nonlinear $P_*(\kappa)$ complementarity problems", Preprint, Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada, December 2000.

[27] Potra, F.A., "On Q-order and R-order of convergence", *Journal of Optimization Theory and Applications*, 63 (1989) 415-431.

[28] Potra, F.A., and Sheng, R., "A large-step infeasible interior point method for the P_*-matrix LCP", *SIAM Journal on Optimization*, 7 (1997) 318-335.

[29] Potra, F.A., and Sheng, R., "Homogeneous interior-point algorithms for semidefinite programming", *Optimization Methods and Software*, 9 (1998) 161-184.

[30] Potra, F.A., and Ye, Y., "Interior point methods for nonlinear complementarity problems", *Journal on Optimization Theory and Applications*, 88 (1996) 617-647.

[31] Sun, J., and Yhao, G., "A quadratically convergent polynomial long-step algorithm for a class of nonlinear complementarity problems", Working paper, National University of Singapore, Republic of Singapore 119260, December 1995.

[32] Tseng, P., "Analysis of an infeasible interior path-following method for complementarity problems", Report, Department of Mathematics, University of Washington, Seattle, Washington, USA, September 1997.

[33] Tseng, P., "An infeasible path-following method for monotone complementarity problems", *SIAM Journal on Optimization*, 7 (1997) 386-402.

[34] Väliaho, H., "P_*-matrices are just sufficient", *Linear Algebra and its Applications*, 239 (1996) 103-108.

[35] Wright, S., and Ralph, D., "Superlinear convergence of an interior-point method for 391 monotone variational inequalities", *Mathematics of Operations Research*, 21 (1996) 815-838.

[36] Ye, Y., "On homogeneous and self-dual algorithm for LCP", *Mathematical Programming*, 76 (1997) 211-222.

[37] Ye, Y., and Anstreicher, K., "On quadratic and $O(\sqrt{n}L)$ convergence of predictor-corrector algorithm for LCP", *Mathematical Programming*, 62 (3) (1993) 537-551.

[38] Ye, Y., Todd, M., and Mizuno, S., "An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm", *Mathematics of Operations Research*, 19 (1994) 53-67.